AD-A102 570 CALIFORNIA UNIV BERKELEY CENTER FOR PURE AND APPLIE--ETC F/6 12/1
METHODS FOR SCALING TO DOUBLY STOCHASTIC FORM,(U)
JUN 81 B N PARLETT, T L LANDIS
UNCLASSIFIED PAM-44 NL

1 OF 1
AD A
102570

END
DATE
FILMED
9-81
DTIC

LEVEL II

AD A102570

CENTER FOR PURE AND APPLIED MATHEMATICS
UNIVERSITY OF CALIFORNIA, BERKELEY

PAM-44

METHODS FOR SCALING TO DOUBLY STOCHASTIC FORM

B. N. Parlett and T. L. Landis

June 26, 1981

Legal Notice

14 PAM-44

Methods for Scaling to Doubly Stochastic Form

B. N. Parlett[1] *and T. L. Landis*[2]

ABSTRACT

New methods for scaling square, nonnegative matrices to doubly stochastic form are described. A generalized version of the convergence theorem in SINKHORN and KNOPP [1967] is proved and applied to show convergence for these new methods. Tests indicate that one of the new methods has significantly better average and worst-case behavior than the Sinkhorn/Knopp method; for one of the 3×3 examples in MARSHALL and OLKIN [1968], SK requires 130 times as many operations as the new algorithm to achieve row and column sums 1 ± 10^{-5}.

1. Introduction.

We seek an algorithm which will find a pair of positive diagonal matrices D and E for a given square nonnegative matrix, A, such that DAE is doubly stochastic--or determine that such a pair does not exist.

A nonnegative $n \times n$ matrix A is said to have *support* if it possesses a positive diagonal; A has *total support* if $A \neq 0$ and every positive entry in A lies on a positive diagonal. A is *fully indecomposable* if it is impossible to find permutation matrices P and Q so that:

$$PAQ = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix} \tag{1.1}$$

with A_1 square.

[1] Department of Mathematics, Computer Science Division of EECS Department, and Electronics Research Laboratory, University of California, Berkeley, CA 94720

[2] Department of Mathematics, University of California, Berkeley, CA 94720

The authors gratefully acknowledge partial support from Office of Naval Research contract N00014-76-C-0013.

Three procedures for computing D and E, when they exist, have appeared in the literature: (1) minimize $x^t A y$ subject to the constraints $\prod x_i = \prod y_i = 1$, (2) minimize

$$f(x_1, \cdots, x_n) = \frac{\prod_{k=1}^{n} \left[\sum_{i=1}^{n} a_{ki} x_i \right]}{\prod_{k=1}^{n} x_k}$$

subject to the constraints

$$x_k > 0 \quad k = 1, \cdots, n \text{ and } \sum_{k=1}^{n} x_k = 1$$

and (3) compute D and E iteratively by alternately normalizing all rows and all columns in A. The first method is due to MARSHALL and OLKIN [1968]; the second is described in DJOKVIC [1970]. In each case, the minimization problem is shown to have a solution when A is fully indecomposable. The third algorithm was first described by DEMING and STEPHAN [1940] who called it the "Iterative Proportional Fitting Procedure". It was rediscovered by SINKHORN [1962, 1964, 1966, 1967], and, SINKHORN and KNOPP [1967] proved that D and E exist so that $D A E$ is doubly stochastic if and only if A possesses total support. Further, they showed that in such a case, the iteration converges to a solution pair D and E. BRUALDI, PARTER, and SCHNEIDER [1966] independently proved the existence of D and E when A is a direct sum of fully indecomposable matrices by showing that its corresponding Menon operator (MENON [1967]) has a fixed point. Finally, SINKHORN [1966] showed that the Sinkhorn/Knopp method converges geometrically for positive starting matrices.

The following result can be applied to show that a nonnegative matrix has total support if and only if it is a direct sum of fully indecomposable matrices:

The Frobenius–Konig Theorem (MARCUS and MINC [1964], p 97)

A nonnegative $n \times n$ matrix without support contains an $s \times t$ zero submatrix where:

$$s + t = n + 1$$

In this paper we describe three new iterative procedures for scaling nonnegative matrices to doubly stochastic form. We prove a generalized version of the convergence theorem in SINKHORN and KNOPP [1967] and apply it to show that for starting matrices with total support, these new iterations converge to diagonally equivalent limits which are multiples of doubly stochastic matrices. In the final--and most interesting--section, we present results of tests comparing our new methods to the Sinkhorn/Knopp method (SK). One of the new algorthms, EQ, exhibited significantly better average and worst-case behavior than SK: for some test matrices, SK required 130 times as maan operations as EQ (where an operation is a multiply or a divide) and examples for which EQ requires more than ten times as many operations as SK are rare.

Techniques for scaling to doubly stochastic form have a number of applications. The problem that launched Sinkhorn's research was estimating the transition matrix in a Markov chain. MARSHALL and OLKIN [1979] contains references for other statistical applications. They can be applied to equillibrate a general matrix with respect to any p-norm, $p \neq \infty$; one of us has used EQ to test for diagonal equivalence to orthogonal form by equillibrating with respect to the 2-norm. Finally, we remark that doubly stochastic matrices possess the following interesting properties: (1) They are "perfectly balanced" with respect to the 1-norm (see PARLETT and REINSCH [1962]); (2) Their p-norms are unity for all $p \leq \infty$ (see STOER and WITZGALL [1962]); and (3) Their inverses--if they exist--have row and column sums equal to unity (though the inverse of a doubly stochastic matrix is doubly stochastic only for permutation matrices).

2. Algorithms

In this section, we will describe three iterative procedures for scaling non-negative matrices to doubly stochastic form. In section three we show that when they are applied to a matrix with total support, the result is a sequence of iteration matrices converging to a multiple of a doubly stochastic matrix.

Our first algorithm, DEV, was motivated by the desire to have an algorithm that would modify a single row or column--leaving the remainder of the matrix unchanged--at each iterative step. There is a natural way to select the row or column to be changed: choose one whose sum deviates maximally from the mean of the row sums (which is also the mean of the column sums). This approach is reasonable, because matrices with equal row and column sums are scalar multiples of doubly stochastic matrices. For the same reason, the natural change is to multiply entries in the selected row or column by a factor chosen so that its new sum will be the new mean of row sums.

Algorithm 1 (called DEV, for deviation reduction)

Given $A = A^{(0)}$ an $n \times n$ matrix:

(1) Compute row and column sums for A:

$$r_i \leftarrow \sum_{j=1}^{n} a_{ij} \quad i = 1,\ldots,n$$

$$c_j \leftarrow \sum_{i=1}^{n} a_{ij} \quad j = 1,\ldots,n$$

Compute the mean, μ, of row sums in A:

$$\mu \leftarrow \frac{1}{n} \left(\sum_{i=1}^{n} r_i \right)$$

(2) Find indices p and q so that

$$| r_p - \mu | = \max_i | r_i - \mu |$$

and

$$| c_q - \mu | = \max_j | c_j - \mu |$$

If $|r_p - \mu| < tol\ \mu$ and $|c_q - \mu| < tol\ \mu$ go to step 5.

If $|c_q - \mu| > |r_p - \mu|$ go to step 4.

(3) Calculate the mean $\bar{\mu}$, of row sums other than r_p :

$$\bar{\mu} \leftarrow \frac{1}{n-1} \left(\sum_{\substack{i=1 \\ i \neq p}}^{n} r_i \right)$$

Scale row p to $\bar{\mu}$:

$$a_{pj} \leftarrow a_{pj} \cdot \frac{\bar{\mu}}{r_p} \quad j = 1, \cdots ,n$$

Update row and column sums for A:

$$r_p \leftarrow \bar{\mu}$$

$$c_j \leftarrow c_j + \left(\frac{\bar{\mu}}{r_p} - 1 \right) a_{pj} \quad j = 1, \cdots ,n$$

Go to step 2.

(4) Calculate the mean, $\bar{\mu}$, of column sums other than c_q:

$$\bar{\mu} \leftarrow \frac{1}{n-1} \left(\sum_{\substack{j=1 \\ j \neq q}}^{n} c_j \right)$$

Scale column q to $\bar{\mu}$:

$$a_{iq} \leftarrow a_{iq} \cdot \frac{\bar{\mu}}{c_q} \quad i = 1, \cdots ,n$$

Update row and column sums for A:

$$c_q \leftarrow \bar{\mu}$$

$$r_i \leftarrow r_i + \left(\frac{\bar{\mu}}{c_q} - 1 \right) \cdot a_{iq}$$

Go to step 2.

(5) Normalize:

$$a_{ij} \leftarrow \frac{1}{\mu} a_{ij} \quad i,j = 1, \cdots ,n$$

Exit.

Remarks

Note that step 3 is equivalent to premultiplying matrix A by a positive diagonal matrix:

$$D = diag\,(d_1, \cdots, d_n)$$

where

$$d_i = \begin{cases} 1 & \text{if } i \neq p \\ \dfrac{\mu}{r_p} & \text{if } i = p \end{cases}$$

and step 4 is equivalent to post multiplying matrix A by a positive diagonal matrix:

$$E = diag\,(e_1, \cdots, e_n)$$

where

$$e_j = \begin{cases} 1 & \text{if } j \neq q \\ \dfrac{\mu}{c_q} & \text{if } j = q \end{cases}$$

We say that a row and column pair in a nonnegative matrix is *balanced* (with respect to the 1-norm) if they have equal sums. Obviously, all row and column sums in a multiple of a doubly stochastic matrix are balanced. A second approach to scaling to doubly stochastic form, then, is to find a row and a column whose sums have maximal difference and to scale the matrix so that their sums are equal. This is the approach taken by our second algorithm.

Algorithm 2 (called BAL, for balance)

Given $A = A^{(0)}$ an $n \times n$ nonnegative matrix:

(1) Compute row and column sums for A:

$$r_i \leftarrow \sum_{j=1}^{n} a_{ij} \quad i = 1, \cdots, n$$

$$c_j \leftarrow \sum_{i=1}^{n} a_{ij} \quad j = 1, \cdots, n$$

and the mean of row sums:

$$\mu \leftarrow \frac{1}{n}\left(\sum_{i=1}^{n} r_i\right)$$

(2) Find indices p and q so that:

$$|r_p - c_q| = \max_{i,j} |r_i - c_j|$$

If $|r_p - c_q| < \mu \cdot tol$ go to step 5

(3) Balance row p and column q:

Multiply entries in row p by:

$$f = \left(\frac{c_q - a_{pq}}{r_p - a_{pq}}\right)^{1/2}$$

and multiply entries in column q by f^{-1}.

(4) Update row and column sums:

$$r_i \leftarrow r_i + (f^{-1} - 1)\, a_{iq} \quad i = 1, \cdots, n$$

$$r_p \leftarrow \left\{(r_p - a_{pq})(c_q - a_{pq})\right\}^{1/2} + a_{pq}$$

$$c_q \leftarrow r_p$$

$$c_j \leftarrow c_j + (f - 1)\, a_{pj} \quad j = 1, \cdots, n$$

$$\mu \leftarrow \frac{1}{n} \sum_{i=1}^{n} r_i$$

Go to step 2.

(5) Normalize:

$$a_{ij} \leftarrow \frac{1}{\mu} a_{ij} \quad i,j = 1, \cdots, n$$

Exit.

Note that step 3 is equivalent to forming the product $D\,A\,E$ where:

$$D = diag(d_1, \cdots, d_n)$$

$$d_i = \begin{matrix} 1, & \text{if } i \neq p \\ f, & \text{if } i = p \end{matrix}$$

$$E = diag(e_1, \cdots, d_n)$$

$e_j = 1, \quad$ if $j \neq q$,

$$= f^{-1}, \quad \text{if } j = q$$

Now for the third method. When testing DEV we found cases where a sequence of 10 or more iterations were alternately scaling the same row and the same column. Our third algorithm is a variant of DEV that avoids this problem. It records the last row and last column scaled; when it detects a repeat, it performs a balancing step.

Algorithm 3 (called EQ, for equalize)

Given $A = A^{(0)}$ an $n \times n$ nonnegative matrix:

(1) Initialize:

$$lastr \leftarrow 0$$

$$lastc \leftarrow 0$$

$$r_i \leftarrow \sum_{j=1}^{n} a_{ij} \quad i = 1, \cdots, n$$

$$c_j \leftarrow \sum_{i=1}^{n} a_{ij} \quad j = 1, \cdots, n$$

$$\mu \leftarrow \frac{1}{n}\left(\sum_{i=1}^{n} r_i\right)$$

(2) Find indices p and q so that:

$$|r_p - \mu| = \max_i |r_i - \mu|$$

$$|c_q - \mu| = \max_j |c_j - \mu|$$

If $|r_p - \mu| < \mu \cdot tol$ and $|c_q - \mu| < \mu \cdot tol$ go to step 6.

If $|r_p - \mu| < |c_q - \mu|$ go to step 4.

(3) If $p = lastr$ go to step 5. Otherwise, calculate the mean, $\bar{\mu}$, of row sums other than r_p:

$$\bar{\mu} \leftarrow \frac{1}{n-1}\left(\sum_{\substack{i=1 \\ i \neq p}}^{n} r_i\right)$$

and scale row p to $\bar{\mu}$:

$$a_{pj} \leftarrow a_{pj} \left(\frac{\bar{\mu}}{r_p} \right) \quad j = 1, \quad ,n$$

Update row and column sums for A:

$$r_p \leftarrow \bar{\mu}$$

$$c_j \leftarrow c_j + \left(\frac{\bar{\mu}}{r_p} - 1 \right) a_{pj} \quad j = 1, \cdots ,n$$

$$\mu \leftarrow \bar{\mu}$$

$$lastr \leftarrow p$$

Go to step 2.

(4) If $q = lastc$ go to step 5. Otherwise, calculate the mean, $\bar{\mu}$, of column sums other than c_q:

$$\bar{\mu} \leftarrow \frac{1}{n-1} \left(\sum_{\substack{j=1 \\ j \neq q}}^{n} c_j \right)$$

and scale column q to $\bar{\mu}$:

$$a_{iq} \leftarrow a_{iq} \cdot \frac{\bar{\mu}}{c_q} \quad i = 1, \cdots ,n$$

Update row and column sums:

$$c_q \leftarrow \bar{\mu}$$

$$r_i \leftarrow r_i + \left(\frac{\bar{\mu}}{c_q} - 1 \right) \quad i = 1, \quad ,n$$

$$\mu \leftarrow \bar{\mu}$$

$$lastc \leftarrow q$$

Go to step 2.

(5) Balance row lastr and column lastc (for convenience let $k = lastr$ and $l = lastc$):

Multiply entries in row k by:

$$f = \left(\frac{c_l - a_{kl}}{r_k - a_{kl}} \right)^{1/2}$$

$$a_{kj} \leftarrow a_{kj}\, f \qquad j = 1, \cdots, n$$

Multiply entries in column l by f^{-1}:

$$a_{il} \leftarrow a_{il}\, f^{-1} \qquad i = 1, \cdots, n$$

Update row and column sums:

$$r_i \leftarrow r_i + (f^{-1} - 1)a_{il} \qquad i = 1, \cdots, n$$

Update row and column sums:

$$r_i \leftarrow r_i + (f^{-1} - 1)\, a_{il} \qquad i = 1, \cdots, n$$

$$r_k \leftarrow \left\{(r_k - a_{kl})(c_l - a_{kl})\right\}^{1/2} + a_{kl}$$

$$c_l \leftarrow r_k$$

$$c_j \leftarrow c_j + (f - 1)a_{kj} \qquad j = 1, \cdots, n$$

$$\mu \leftarrow \frac{1}{n}\left\{\sum_{i=1}^{n} r_i\right\}$$

Go to 2.

(6) Normalize:

$$a_{ij} \leftarrow \frac{1}{\mu} a_{ij} \qquad i,j = 1, \cdots, n$$

Exit.

Notation

To simplify the descriptions of the algorithms we have omitted programming details. In particular, we have assumed that all scaling and balancing operations are carried out explicitly by modifying entries in matrix A. In the next section, it will be convenient to assume that the iterations are carried out implicitly by changing entries in a pair of diagonal matrices D and E.

Each algorithm produces a sequence of iteration matrices which are diagonally equivalent to the starting matrix $A = A^{(0)}$:

(2.1) $$A^{(k)} = D^{(k)} A\, E^{(k)} \qquad k = 1, 2, \cdots$$

$$D^{(k)} = diag\,(d_1^{(k)}, \cdots, d_n^{(k)})$$

$$E^{(k)} = diag\,(e_1^{(k)}, \quad , e_n^{(k)})$$

and we set $D^{(0)} = E^{(0)} = I$.

We introduce the following notation:

$$A^{(k)} = \left[a_{ij}^{(k)}\right]$$

(So, for example, $a_{ij}^{(k)} = d_i^{(k)} a_{ij} e_j^{(k)}$)

$$r_i^{(k)} = \sum_{j=1}^{n} a_{ij}^{(k)}$$

$$c_j^{(k)} = \sum_{i=1}^{n} a_{ij}^{(k)}$$

and

$$\mu_k = \frac{1}{n}\left[\sum_{i=1}^{n} r_i^{(k)}\right]$$

3. Convergence

In this section, we will prove that when a starting matrix has total support each of the algorithms described in section 2 produces a sequence of iteration matrices which converges to a diagonally equivalent, doubly stochastic limit.

SINKHORN and KNOPP [1967] showed that when SK is applied to an $n \times n$ nonnegative starting matrix $A^{(0)} = A$ possessing nonzero row and column sums, the result is a sequence of iteration matrices as in (2.1) with the following properties:

(P1) The sequence $(s_k)_{k=1,2,\ldots}$ is monotonically increasing where:

(3.1) $$s_k = \prod_{i=1}^{n} d_i^{(k)} e_i^{(k)} \qquad k = 1,2,\ldots$$

(P2) If $\lim\limits_{k \to \infty} \dfrac{s_k}{s_{k+1}} = 1$ then for $i,j = 1, \ldots ,n$:

$$\lim_{k \to \infty} r_i^{(k)} = 1$$

$$\lim_{k \to \infty} \frac{d_i^{(k+1)}}{d_i^{(k)}} = 1$$

$$\lim_{k \to \infty} c_j^{(k)} = 1$$

$$\lim_{k \to \infty} \frac{e_j^{(k+1)}}{e_j^{(k)}} = 1$$

(P3) With the k^{th} mean of row sums, μ_k, defined by (2.3)

$$\mu_k = 1 \qquad k = 1,2,\ldots$$

Algorithms which, given an $n \times n$ nonnegative starting matrix, A, produce a sequence of iteration matrices as in (2.1) satsifying (P1), (P2), and (P3), will be called "diagonal product increasing" (DPI). The following result is a simple generalization of the convergence theorem in SINKHORN and KNOPP [1967].

Theorem 1

Given a sequence (2.1) of diagonal equivalents for A satisfying (P1), (P2), and (P3):

[1] If A has support then $\lim_{k\to\infty} A^{(k)}$ exists and is doubly stochastic.

[2] If A has total support then the limit in [1] is diagonally equivalent to A.

Before proving the theorem we state and prove a corollary:

Corollary

[1] If A is diagonally equivalent to a doubly stochastic matrix, S, then:

$$S = \lim_{k\to\infty} A^{(k)}$$

[2] If A has support and is not diagonally equivalent to a doubly stochastic matrix, then for each pair of indices (i,j) such that a_{ij} does not lie on a positive diagonal:

$$\lim_{k\to\infty} a_{ij}^{(k)} = 0$$

Proof of Corollary:

(1) By Birkhoff's Theorem (BIRKHOFF [1946]) the set of $n \times n$ doubly stochastic matrices is the convex hull of the set of $n \times n$ permutation matrices. Therefore, S and its diagonal equivalent, A, have total support. Now the theorem implies that $\lim_{k\to\infty} A^{(k)}$ is doubly stochastic and diagonally equivalent to A. Since doubly stochastic equivalents are unique, (see SINKHORN and KNOPP [1967]):

$$\lim_{k\to\infty} A^{(k)} = S$$

(2) By the theorem, $\lim_{k\to\infty} A^{(k)}$ is doubly stochastic, so it has total support, and $\lim_{k\to\infty} a_{ij}^{(k)} = 0$ whenever a_{ij} does not lie on a positive diagonal. ■

Note that matrices without support are not covered by the preceding theorem or its corollary. Such matrices are always singular, and KAHAN [private communication] has shown that the sequence of iteration matrices ($A^{(k)}$) produced by SK cycles for such a starting matrix.

Proof of Theorem 1:

We will need the following well known result:
Lemma 1 (*The Arithmetic / Geometric Mean Inequality*)

If $x_i \ge 0$ for $i = 1, \cdots, n$ then:

$$\prod_{i=1}^{n} x_i \le \left[\frac{\sum_{i=1}^{n} x_i}{n} \right]^n$$

with equality only when $x_1 = x_2 = \cdots = x_n$

(1): (P1) implies $(s_k)_{k=1,2,\ldots}$ is monotonically increasing. Since A has support, a permutation, σ, of $\{1, \cdots, n\}$ exists such that:

$$\left\{ a_{i,\sigma(i)} \mid i = 1, \cdots, n \right\}$$

is a positive diagonal in A. Let $a = \min_i (a_{i,\sigma(i)})$. Then

$$\sum_{i=1}^{n} d_i^{(k)} e_{\sigma(i)}^{(k)} a \le \sum_{i=1}^{n} d_i^{(k)} e_{\sigma(i)}^{(k)} a_{i,\sigma(i)} = \sum_{i=1}^{n} a_{i\sigma(i)}^{(k)} \le \sum_{i=1}^{n} r_i^{(k)} = n$$

(Property (P3) is used for the right hand equality) By the arithmetic/geometric inequality:

$$s_k = \prod_{i=1}^{n} d_i^{(k)} e_{\sigma(i)}^{(k)} \le a^{-n}$$

and $\left\{s_k\right\}_{k=1,2,\ldots}$ is bounded. Therefore by (P1)

$$\lim_{k \to \infty} s_k = L > 0$$

exists, and

$$\lim_{k \to \infty} \frac{s_k}{s_{k+1}} = 1$$

By (P2):

$$\lim_{k \to \infty} \frac{d_i^{(k+1)}}{d_i^{(k)}} = 1 \quad \text{and} \quad \lim_{k \to \infty} \frac{e_j^{(k+1)}}{e_j^{(k)}} = 1$$

By (P3), since the $A^{(k)}$ are nonnegative, no entry $a_{ij}^{(k)}$ can be larger than n. Therefore, for each index pair (i,j) the sequence $(a_{ij}^{(k)})$ is Cauchy, and

$$\lim_{k\to\infty} A^{(k)} = A^{(\infty)}$$

exists. Since the row and column sums in $A^{(\infty)}$ must be:

$$\lim_{k\to\infty} r_i^{(k)} = r_i^{(\infty)} \qquad i = 1, \cdots, n$$

$$\lim_{k\to\infty} c_j^{(k)} = c_j^{(\infty)} \qquad j = 1, \cdots, n$$

(P2) implies that $A^{(\infty)}$ is doubly stochastic.

(2): To prove the second half of the theorem we will need the following lemma which is paraphrased from SINKHORN and KNOPP [1967], p. 345.

Lemma 2

If A is a nonnegative matrix with total support, $(x_i^{(k)})$ and $(y_i^{(k)})$ are positive sequences for $i = 1, \cdots, n$ and $j = 1, \cdots, n$ and:

$$\lim_{k\to\infty} x_i^{(k)} y_j^{(k)} = L_{ij} > 0$$

for each index pair (i,j) such that $a_{ij} \neq 0$, then there exist positive sequences $(\bar{x}_i^{(k)})$ and $(\bar{y}_j^{(k)})$ with positive limits such that:

$$\bar{x}_i^{(k)} \bar{y}_j^{(k)} = x_i^{(k)} y_j^{(k)} \quad \text{for all } i, j, \text{ and } k$$

Now for the proof. From part (1), we know that $\lim_{k\to\infty} a_{ij}^{(k)} = \lim_{k\to\infty} d_i^{(k)} e_j^{(k)} a_{ij}$ exists for any i and j. If $a_{ij} \neq 0$ then $\lim_{k\to\infty} d_i^{(k)} e_j^{(k)}$ exists. Using (P1) we show that this limit is positive.

If $a_{ij} \neq 0$, it lies on a positive diagonal in A, because A has total support. Let σ be a permutation of $\{1,\ldots,n\}$ such that:

$$\sigma(i) = j$$

$$a_{l\sigma(l)} > 0 \qquad l = 1, \cdots, n$$

By (P1):

$$d_i^{(k)} e_j^{(k)} \prod_{\substack{l=1 \\ l\neq i}}^{n} d_l^{(k)} e_{\sigma(l)}^{(k)} = s_k \geq s_1 \qquad k = 1, 2, \cdots$$

(3.2) $$d_i^{(k)} e_j^{(k)} \geq s_1 \left(\prod_{\substack{l=1 \\ l\neq i}} d_l^{(k)} e_{\sigma(l)}^{(k)} \right)^{-1} \qquad k = 1, 2, \cdots$$

Let $a = \min_{i,j} \left\{ a_{ij} \mid a_{ij} \neq 0 \right\}$. Then:

$$a \left(\sum_{\substack{l=1 \\ l \neq i}}^{n} d_l^{(k)} e_{\sigma(l)}^{(k)} \right) \leq \sum_{\substack{l=1 \\ l \neq i}}^{n} d_l^{(k)} e_{\sigma(l)}^{(k)} a_{l,\sigma(l)} \leq \sum_{i=1}^{n} r_i^{(k)} = n$$

$$\frac{a \left(\sum_{\substack{l=1 \\ l \neq i}}^{n} d_l^{(k)} e_{\sigma(l)}^{(k)} \right)}{n-1} \leq \frac{n}{n-1}$$

Now apply the arithmetic/geometric mean inequality:

$$\prod_{\substack{l=1 \\ l \neq i}}^{n} d_l^{(k)} e_{\sigma(l)}^{(k)} \leq \left(\frac{n \cdot a^{-1}}{n-1} \right)^{n-1} ,$$

or

$$(3.3) \qquad \left(\prod_{\substack{l=1 \\ l \neq i}}^{n} d_l^{(k)} e_{\sigma(l)}^{(k)} \right)^{-1} \geq \left(\frac{(n-1)\, a}{n} \right)^{n-1}$$

Combining (3.2) and (3.3):

$$(3.4) \qquad d_i^{(k)} e_j^{(k)} \geq s_1 \left(\frac{(n-1)\, a}{n} \right)^{n-1} > 0$$

which shows that $\lim_{k \to \infty} d_i^{(k)} e_j^{(k)} > 0$ whenever $a_{ij} \neq 0$. Now we can apply lemma 2 to see that positive sequences $\left(\bar{d}_i^{(k)}\right)$ and $\left(\bar{e}_j^{(k)}\right)$ with positive limits exist so that :

$$\bar{d}_i^{(k)} \bar{e}_j^{(k)} = d_i^{(k)} e_j^{(k)} \qquad \text{for } each \ i, j, \text{ and } k$$

Set

$$\bar{D}^{(k)} = diag\,(\bar{d}_i^{(k)})$$

and

$$\bar{E}^{(k)} = diag\,(\bar{e}_i^{(k)})$$

then

$$\lim_{k \to \infty} \bar{D}^{(k)} = \bar{D}^{(\infty)} \quad \text{and}$$

$$\lim_{k \to \infty} \bar{E}^{(k)} = \bar{E}^{(\infty)}$$

exist. Taking limits on both sides of

$$\bar{D}^{(k)} A \, \bar{E}^{(k)} = A^{(k)}$$

we obtain

$$D^{(\infty)} A \, E^{(\infty)} = A^{(\infty)} \qquad \blacksquare$$

Naturally, the Sinkhorn/Knopp method is product increasing; in the next theorem, we will show that normalized versions of DEV, BAL, and EQ, are too. Here is another example of a DPI algorithm defined for irreducible, nonnegative matrices:

Algorithm

At each step: Normalize the rows by finding Y, a positive diagonal matrix, so that $YA^{(k)}$ has row sums 1. Then normalize the columns by a diagonal similarity transform defined as follows:

Let $x = (x_1, \cdots, x_n)$ be a left Perron vector for $YA^{(k)}$:

$$x \, Y A^{(k)} = 1 \cdot x$$

and let $X = diag\,(x_1, \cdots, x_n)$. Then

$$A^{(k+1)} = \left\{X \, Y\right\} A^{(k)} X^{-1}$$

has column sums 1 because

$$(1, \cdots, 1) A^{(k+1)} = (1, \cdots, 1)$$

(Note that the similarity transform leaves diagonal products unchanged)

Next we apply Theorem 1 to show that the algorithms described in section 2 are convergent for starting matrices with total support.

Theorem 2

Suppose that the sequence of iteration matrices

$$A^{(k)} = D^{(k)} A \, E^{(k)} \quad k = 1, \cdots, n$$

results from the application of DEV, BAL, or EQ, to $A = A^{(0)}$; then if A has total support, $\lim_{k \to \infty} \frac{A^{(k)}}{\mu_k}$ is doubly stochastic and diagonally equivalent to A.

Proof

We prove Theorem 2 by showing that the sequence of normalized iteration matrices:

$$\bar{A}^{(k)} = \frac{A^{(k)}}{\mu_k} = \bar{D}^{(k)} A\, E^{(k)} = \left(\frac{1}{\mu_k} D^{(k)}\right) A\, E^{(k)} \qquad k = 1,2, \cdots \tag{3.5}$$

satisfies (P1), (P2), and (P3).

(P3) is obviously satisfied by (3.5). Note that:

$$s_k = \prod_{i=1}^{n} \left(\frac{1}{\mu_k} d_i^{(k)}\right) e_i^{(k)}$$

$$= \frac{1}{\mu_k^n} \prod_{i=1}^{n} d_i^{(k)} e_i^{(k)} \qquad k = 1,2, \cdots$$

DEV:

Suppose that at step k+1 row p is scaled to the mean of the other row sums, $\bar{\mu}$. After the scaling, $\bar{\mu}$ is the mean of row sums, that is:

$$\bar{\mu} = \mu_{k+1}$$

and in this case:

$$\frac{s_{k+1}}{s_k} = \left(\frac{1 / \mu_{k+1}}{1 / \mu_k}\right)^n \cdot \frac{\prod_{i=1}^{n} d_i^{(k+1)} e_i^{(k+1)}}{\prod_{i=1}^{n} d_i^{(k)} e_i^{(k)}} \tag{3.6}$$

$$= \left(\frac{1 / \mu_{k+1}}{1 / \mu_k}\right)^n \frac{\mu_{k+1}}{r_p}$$

$$= \left(\frac{\mu_k}{\mu_{k+1}}\right)^{n-1} \cdot \frac{\mu_k}{r_p}$$

Since μ_{k+1} is the mean of row sums other than r_p in $A^{(k)}$, and μ_k is the mean of all row sums:

$$\frac{r_p + (n-1)\mu_{k+1}}{n} = \mu_k .$$

$$\frac{1}{n}\left(\frac{r_p}{\mu_k} + \frac{(n-1)\mu_{k+1}}{\mu_k}\right) = 1 \tag{3.7}$$

By the arithmetic/geometric mean inequality:

$$\frac{r_p \cdot \mu_{k+1}^{n-1}}{\mu_k^n} \leq 1$$

Therefore:

$$\frac{s_{k+1}}{s_k} = \left(\frac{\mu_k}{\mu_{k+1}}\right)^{n-1} \cdot \frac{\mu_k}{r_p} \geq 1$$

The above arguement can be repeated for a column scaling at step (k+1), and (P1) holds. Next define sequences:

(3.8) $$\left\{x_i^{(k)}\right\}_{k=1,2,\cdots} \qquad i = 1, \cdots, n$$

by

$$x_i^{(k+1)} = \begin{cases} \dfrac{r_p}{\mu_k} & \text{if } i=p \text{ and at step (k+1) row } p \text{ is scaled to the mean of the other row sums} \\ \dfrac{c_q}{\mu_k} & \text{if } i=q \text{ and at step (k+1) column q is scaled to the mean of the other column sums} \\ \dfrac{\mu_{k+1}}{\mu_k} & \text{otherwise} \end{cases}$$

By (3.7), $\frac{1}{n}\sum_{i=1}^{n} x_i^{(k)} = 1$ for each k. Using the arithmetic/geometric mean inequality it can be shown that from:

$$\lim_{k \to \infty} \prod x_i^{(k+1)} = \lim_{k \to \infty} \frac{s_{k+1}}{s_k} = 1$$

follows:

$$\lim_{k \to \infty} x_i^{(k)} = 1 \quad i = 1, \cdots, n$$

Since

$$\frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \frac{\mu_k}{\mu_{k+1}} \quad or \quad \frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \frac{\mu_{k+1}}{r_p} \cdot \frac{\mu_k}{\mu_{k+1}} = \frac{\mu_k}{r_p} \qquad i = 1, \cdots, n$$

$$\frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \frac{1}{x_l^{(k+1)}} \quad \text{for } some \ l$$

and similarly

$$\frac{e_i^{(k+1)}}{e_i^{(k)}} = 1 \ \ or \ \ \frac{\mu_k}{c_q} \qquad i = 1, \cdots, n$$

$$= 1 \ \ or \ \ \frac{1}{x_l^{(k+1)}} \quad \text{for } some \ l$$

it follows that:

$$\lim_{k \to \infty} \frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \lim_{k \to \infty} \frac{e_i^{(k+1)}}{e_i^{(k)}} = 1 \qquad i = 1, \cdots, n$$

At each step, DEV selects p or q so that

$$\left| \frac{r_p}{\mu_k} - 1 \right| \ \ or \ \ \left| \frac{c_q}{\mu_k} - 1 \right|$$

is maximal. It follows that for $i,j = 1, \cdots, n$:

$$\lim_{k \to \infty} \frac{r_i^{(k)}}{\mu_k} = \lim_{k \to \infty} \frac{c_j^{(k)}}{\mu_k} = 1$$

Therefore (P2) holds for the sequence (3.5) produced by DEV.

BAL:

Suppose that at step (k+1) BAL balances row p and column q. Let:

$$x^{(k)} = r_p^{(k)} - a_{pq}^{(k)}$$

$$y^{(k)} = c_q^{(k)} - a_{pq}^{(k)}$$

In this case:

(3.9) $$\frac{s_{k+1}}{s_k} = \left[\frac{1 \ / \ \mu_{k+1}}{1 \ / \ \mu_k} \right]^n f \quad f^{-1} = \left[\frac{\mu_k}{\mu_{k+1}} \right]^n$$

where

$$f = \left[\frac{x^{(k)}}{y^{(k)}} \right]^{1/2}$$

so to show that $s_{k+1} \geq s_k$, we must show that $\mu_{k+1} \leq \mu_k$.

If $A^{(k)}$ is doubly stochastic for some k, then $A^{(k)} = A^{(k+1)} = \cdots$ and the theorem is satisfied. Otherwise, for each k, we may assume that

$$(3.10) \qquad |r_p^{(k)} - c_q^{(k)}| = |x^{(k)} - y^{(k)}| \neq 0$$

With this assumption, if $x^{(k)} = 0$ or $y^{(k)} = 0$ then A cannot have total support--a contradiction. Therefore, $x^{(k)} \neq 0$ and $y^{(k)} \neq 0$, and the denominators for f and f^{-1} are never 0.

Entries in $A^{(k)}$ sum to:

$$(3.11) \qquad \sum_{i=1}^{n} r_i^{(k)} = \sum_{i,j=1}^{n} a_{ij}^{(k)}$$

$$= \left(\sum_{\substack{i,j=1 \\ i \neq p \\ j \neq q}}^{n} a_{ij}^{(k)} \right) + x^{(k)} + y^{(k)} + 2a_{pq}^{(k)}$$

After the balancing step, entries in $A^{(k+1)}$ sum to:

$$(3.12) \qquad \sum_{i=1}^{n} r_i^{(k+1)} = \left(\sum_{\substack{i=1 \\ i \neq p \\ i \neq q}}^{n} a_{ij}^{(k)} \right) + 2\left(x^{(k)} y^{(k)}\right)^{1/2} + 2a_{pq}^{(k)}$$

so

$$(3.13) \qquad \mu_{k+1} = \frac{\sum_{i=1}^{n} r_i^{(k+1)}}{n}$$

$$= \frac{\sum_{i=1}^{n} r_i^{(k)} + \left[2\sqrt{x^{(k)} y^{(k)}} - \left(x^{(k)} + y^{(k)}\right)\right]}{n}$$

$$= \mu_k + \frac{2\sqrt{x^{(k)} y^{(k)}} - \left(x^{(k)} + y^{(k)}\right)}{n}$$

By the arithmetic/geometric mean inequality

$$2\sqrt{x^{(k)} y^{(k)}} \leq x^{(k)} + y^{(k)}$$

and therefore

$$(3.14) \qquad \mu_{k+1} \leq \mu_k$$

By (3.9), $\frac{s_{k+1}}{s_k} = \left(\frac{\mu_k}{\mu_{k+1}}\right)^n \geq 1$ and $s_{k+1} \geq s_k$, ie. (P1) is satisfied.

Suppose that:

(3.15) $$\lim_{k\to\infty} \frac{s_{k+1}}{s_k} = 1$$

then by (3.9) and (3.13):

(3.16) $$1 = \lim_{k\to\infty} \frac{\dfrac{\mu_k}{\mu_k + 2\sqrt{x^{(k)}y^{(k)}} - \left[x^{(k)} + y^{(k)}\right]}}{n}$$

$$= \lim_{k\to\infty} \frac{\mu_k}{\dfrac{(n\ \mu_k) + 2\sqrt{x^{(k)}y^{(k)}} - (x^{(k)} + y^{(k)})}{n}}$$

$$= \lim_{k\to\infty} \frac{1}{1 + \dfrac{2\sqrt{x^{(k)}y^{(k)}} - (x^{(k)} + y^{(k)})}{n\cdot\mu_k}}$$

which implies that

$$\lim_{k\to\infty} \frac{\left[2\sqrt{x^{(k)}y^{(k)}} - (x^{(k)} + y^{(k)})\right]}{n\cdot\mu_k} = 0$$

$$\lim_{k\to\infty}\left[2\sqrt{\frac{x^{(k)}}{\mu_k}\frac{y^{(k)}}{\mu_k}} - \left(\frac{x^{(k)}}{\mu_k} + \frac{y^{(k)}}{\mu_k}\right)\right] = 0$$

It follows from the arithmetic/geometric mean inequality, that this is impossible unless:

$$\lim_{k\to\infty} \frac{x^{(k)}}{\mu_k} - \frac{y^{(k)}}{\mu_k} = 0$$

and

$$\lim_{k\to\infty}\left\{\max_{i,j} \left|\frac{r_i^{(k)}}{\mu_k} - \frac{c_j^{(k)}}{\mu_k}\right|\right\} = 0$$

The mean of the row sums and the mean of the column sums in $\frac{A^{(k)}}{\mu_k}$ is 1, implying:

$$\lim_{k\to\infty} \frac{r_i^{(k)}}{\mu_k} = \lim_{k\to\infty} \frac{c_j^{(k)}}{\mu_k} = 1$$

Equation (3.4) in the proof of Theorem 1 holds whenever (P1) and (P3) are satisfied, and for each index pair (i,j) such that $a_{ij} \neq 0$, the sequence

$$\left\{\frac{a_{ij}^{(k)}}{\mu_k}\right\}_{k=1,2,}$$

is bounded away from zero. Therefore, the sequences $\left\{\frac{x^{(k)}}{\mu_k}\right\}$ and $\left\{\frac{y^{(k)}}{\mu_k}\right\}$ are bounded away from zero, because $x^{(k)} \neq 0$ and $y^{(k)} \neq 0$ for each k. We have:

$$\text{(3.17)} \qquad \lim_{k\to\infty} \frac{x^{(k)}/\mu_k}{y^{(k)}/\mu_k} = \lim_{k\to\infty} \frac{x^{(k)}}{y^{(k)}} = 1$$

Finally, for each i,j, and k:

$$\frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \frac{\mu_k}{\mu_{k+1}} \frac{d_i^{(k+1)}}{d_i^{(k)}} = \frac{\mu_k}{\mu_{k+1}} \quad or \quad \frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \frac{\mu_k}{\mu_{k+1}} \left\{\frac{y^{(k)}}{x^{(k)}}\right\}^{1/2}$$

$$\frac{e_j^{(k+1)}}{e_j^{(k)}} = 1 \quad or \quad \frac{e_j^{(k+1)}}{e_j^{(k)}} = \left\{\frac{x^{(k)}}{y^{(k)}}\right\}^{1/2}$$

Therefore, by (3.15) and (3.17):

$$\text{(3.18)} \qquad \lim_{k\to\infty} \frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \lim_{k\to\infty} \frac{e_j^{(k+1)}}{e_j^{(k)}} = 1$$

for each i and j, and (P2) is satisfied.

EQ:

Each step of EQ is a step of DEV or a balancing step. The arguements above for DEV and BAL show that for each k, $s_{k+1} \geq s_k$, ie. that (P1) holds. Consider the sequence (3.8), and its subsequence

$$\left\{x_i^{(k')}\right\}_{k'=p_1,p_2,}$$

where at steps $k' = p_1, p_2, \cdots$ EQ scaled a row or column to the mean of the other row or column sums. This sequence must be infinite--because lastr and lastc are set to 0 after each balancing step--and the arguement for DEV can be repeated to show that:

(3.19) for $i,j = 1, \cdots, n$

$$\lim_{k'\to\infty} \frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k')}} = 1 \quad i = 1, \cdots, n$$

$$\lim_{k'\to\infty} \frac{e_i^{(k'+1)}}{e_i^{k'}} = 1$$

$$\lim_{k'\to\infty} \frac{r_i^{k'}}{\mu_{k'}} = 1$$

$$\lim_{k'\to\infty} \frac{c_j^{k'}}{\mu_{k'}} = 1$$

Next consider the subsequence

$$\left\{x_i^{(k'')}\right\}_{k''=q_1,q_2,}$$

where at steps $k'' = q_1, q_2, \cdots$ EQ balanced a row and column. If this sequence is infinite, the arguements for BAL (Equations (3.15)-(3.18)) can be repeated to show that when $\lim_{k\to\infty} \frac{s_{k+1}}{s_k} = 1$:

$$\lim_{k''\to\infty} \frac{\bar{d}_i^{(k''+1)}}{\bar{d}_i^{(k'')}} = 1 \tag{3.20}$$

$$\lim_{k''\to\infty} \frac{e_i^{(k'')}}{e_i^{(k'')}} = 1$$

(3.19) and (3.20) imply that

$$\lim_{k\to\infty} \frac{\bar{d}_i^{(k+1)}}{\bar{d}_i^{(k)}} = \lim_{k\to\infty} \frac{e_i^{(k+1)}}{e_i^{(k)}} = 1$$

In particular, $\lim_{k\to\infty} A^{(k)}$ exists, and its row and column sums are:

$$\lim_{k\to\infty} \frac{r_i^{(k)}}{\mu_k} = \lim_{k'\to\infty} \frac{r_i^{(k')}}{\mu_{k'}} = 1$$

$$\lim_{k\to\infty} \frac{c_j^{(k)}}{\mu_{(k)}} = \lim_{k'\to\infty} \frac{c_j^{(k')}}{\mu_{k'}} = 1$$

(P3) holds for the sequence (3.5) produced by EQ. ∎

It is possible to show that each of the sequences $\left\{d_i^{(k)}\right\}$ and $\left\{e_j^{(k)}\right\}$, $i, j = 1, \cdots, n$, produced by SK and BAL are Cauchy. We believe the same to be true for DEV and EQ but are unable to prove it.

4. Test Results

We ran comparison tests of the algorithms described in this paper and the Sinkhorn/Knopp method on a collection of 50 10×10 or smaller matrices. These tests were run on a VAX 11/780 at UC, Berkeley, with 7 significant digits in single precision and 16 digits in double precision. Sums were accumulated in double precision.

For convergence to "tol" accuracy, we required that all row and column sums deviate from the mean, μ, by less than $tol \cdot \mu$. So, in the normalized matrices, row and column sums could not deviate from 1 by more than tol.

The examples in this section were selected to illustrate the following points:

(1) EQ exhibited significantly better average and worst-case behavior than SK. On our test bed, for convergence to $tol = 10^{-5}$, the ratio of total SK operations to total EQ operations varied from a low of 1/2 to a high of more than 130.

(2) We found striking examples where EQ was significantly faster than DEV, BAL, or SK. Since each iteration by EQ scaled a row or column--like DEV--or balanced a row and column pair--like BAL--there is evidence that some mechanism is at work which enables EQ to choose the right operation at the right time.

To facilitate comparing DEV, BAL, and EQ, we counted their "steps" in the following way: each scaling of a row or column counted as 1 step, and each balancing of a row/column pair counted as two steps. In this way, the operation cost (where an operation is a multiplication or division) was the same for each step of each of the three algorithms.

To facilitate comparing EQ and SK we computed the approximate ratio of total operations performed by EQ to total operations performed by SK.

These first four examples were the test matrices in MARSHALL and OLKIN [1968]:

$$A = \begin{pmatrix} 10^4 & 10^2 & 10^2 \\ 10^2 & 1 & 1 \\ 10^2 & 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 10^2 & 1 & 0 \\ 10^2 & 10^3 & 1 \\ 0 & 10^2 & 10^2 \end{pmatrix}$$

$$C = \begin{pmatrix} 10^2 & 10^2 & 0 \\ 10^2 & 10^4 & 1 \\ 0 & 1 & 10^2 \end{pmatrix} \qquad D = \begin{pmatrix} 10^4 & 1 & 0 \\ 10^4 & 10^6 & 1 \\ 0 & 10^4 & 10^4 \end{pmatrix}$$

STEPS TO CONVERGENCE FOR MATRIX A

TOL	DEV	BAL	EQ	SK	$\left[\frac{SK\ OPERATIONS}{EQ\ OPERATIONS}\right]$
10^{-2}	2	18	2	1	.9
10^{-3}	2	20	2	1	.9
10^{-4}	2	26	2	1	.9
10^{-5}	2	32	2	1	.9

STEPS TO CONVERGENCE FOR MATRIX B

TOL	DEV	BAL	EQ	SK	$\left[\frac{SK\ OPERATIONS}{EQ\ OPERATIONS}\right]$
10^{-2}	102	18	21	38	3.3
10^{-3}	201	24	26	75	5.3
10^{-4}	302	34	37	113	5.6
10^{-5}	402	34	46	150	6.0

STEPS TO CONVERGENCE FOR MATRIX C

TOL	DEV	BAL	EQ	SK	$\left(\frac{SK\ OPERATIONS}{EQ\ OPERATIONS}\right)$
10^{-2}	24	18	11	10	1.7
10^{-3}	80	24	19	307	29.8
10^{-4}	1656	30	38	1092	53.1
10^{-5}	3281	34	49	1899	71.5

STEPS TO CONVERGENCE FOR MATRIX D

TOL	DEV	BAL	EQ	SK	$\left(\frac{SK\ OPERATIONS}{EQ\ OPERATIONS}\right)$
10^{-2}	231	22	21	84	7.4
10^{-3}	1895	26	32	706	40.7
10^{-4}	4891	32	34	1830	99.4
10^{-5}	7961	40	40	2983	137.7

Our next examples, R and S, are (5×5) matrices. They illustrate the curious fact mentioned in (2) above.

$$R = \begin{bmatrix} 100 & 1 & 0 & 0 & 0 \\ 0 & 200 & 1 & 0 & 0 \\ 0 & 0 & 300 & 1 & 0 \\ 0 & 0 & 0 & 400 & 1 \\ 1 & 0 & 0 & 0 & 500 \end{bmatrix}$$

STEPS TO CONVERGENCE FOR MATRIX R

TOL	DEV	BAL	EQ	SK	$\left(\frac{SK\ OPERATIONS}{EQ\ OPERATIONS}\right)$
10^{-2}	10	14	9	1	.4
10^{-3}	115	28	31	214	24.4
10^{-4}	1682	2474	51	630	43.6
10^{-5}	3912	4086	68	1067	55.4

$$S = \begin{pmatrix} 40 & 0 & 1 & 1 & 1 \\ 1 & 80 & 0 & 1 & 1 \\ 1 & 1 & 120 & 0 & 1 \\ 1 & 1 & 1 & 160 & 0 \\ 0 & 1 & 1 & 1 & 200 \end{pmatrix}$$

STEPS TO CONVERGENCE FOR MATRIX S

TOL	DEV	BAL	EQ	SK	$\left[\frac{SK\ OPERATIONS}{EQ\ OPERATIONS}\right]$
10^{-2}	39	20	16	15	3.3
10^{-3}	206	100	29	53	6.5
10^{-4}	384	108	47	94	7.1
10^{-5}	570	198	62	136	7.7

Our final five examples are (10×10) upper Hessenberg matrices:

$$H_1 = (h_{ij}) \ \ where \ \ h_{ij} = \begin{cases} 0 \text{ if } j < i-1 \\ 1 \ \ otherwise \end{cases}$$

H_2, H_3, and H_4, each differ from H_1 in a single entry:

the (1,1) entry in H_2 is 100

the (1,2) entry in H_3 is 100

the (1,3) entry in H_4 is 100

H_5 is the result of replacing all diagonal entries in H_1 by 100.

Here is a summary of the results for $tol = 10^{-5}$:

STEPS TO CONVERGENCE FOR MATRICES H_i, $i = 1, \cdots, 5$

MATRIX	DEV	BAL	EQ	SK	$\left[\frac{SK\ OPERATIONS}{EQ\ OPERATIONS}\right]$
H_1	812	748	812	55	.6
H_2	873	926	717	72	.8
H_3	925	952	775	71	.7
H_4	953	948	921	71	.6
H_5	14476	17456	917	1004	8.9

References

BELLMAN, R.: Introduction to Matrix Analysis, McGraw-Hill Book Company [1970]

BIRKHOFF, G.: Tres observaciones sobre le algebra lineal, Rev. univ. nec. Tucuman, ser A, , 147-151, [1946]

BRUALDI, R.A., S.V. PARTER, and H. SCHNEIDER: The diagonal equivalence of a nonnegative matrix to a stochastic matrix, J. Math. Appl., , 31-50 [1966]

DEMING, N.E. and STEPHAN,F.F.: On a least squares adjustment of a sampled frequency table when the expected marginal totals are known, Ann. Math. Statist., 11 ,427-444 [1940]

DJOKVIC, D.Z.: Note on nonnegative matrices, Proc. Amer. Math. Soc., 25, 80-2, [1970]

FIENBERG, S.E.: An iterative procedure for estimation in contingency tables, Ann. Math. Statist.,
41 , 907-917 [1970]

IRELAND, C.T. and S. KULLBACK: Contingency tables with given marginals. Biometrika, 55 , 179-188, [1968]

MARSHALL, A.W. and I.OLKIN: Scaling of matrices to achieve specified row and column sums, Numer. Math., 12 , 83-90 [1968]

-- -- : Inequalities: Theory of Majorization and Its Applications, Academic Press [1979]

MENON, M.V.: Reduction of a matrix with positive elements to a doubly stochastic matrix, Proc. Amer. Math. Soc., 18,244- 247, [1967]

PARLETT, B.N. and C. REINSCH: Balancing a matrix for calculation of eigenvalues and eigenvectors, Numer. Math.,13,293-304, [1969]

SINKHORN, R.: On best doubly stochastic estimates, Notices Am. Soc., , 294, #592-34, [1962]

-- : A relationship between arbitrary positive matrices , Canad. J. Math., ,303-306 [1966]

-- : A relationship between arbitrary positive matrices and stochastic matrices, Ann. Math. Statistics, 35, 876-879, [1964]

— : Diagonal equivalence to matrices with prescribed row and column sums. Am. Math. Monthly, 74, 402-405, [1967]

— and P. KNOPP: Concerning nonnegative matrices and doubly stochastic matrices, Pac. J. Math., 21, 343-348, [1967]

— —: Problems concerning diagonal products in nonnegative matrices, Trans. of the Amer. Math. Soc. , 136, 67-75, [1969]

Appendix: Sample Iteration matrices for EQ and SK

The following pages contain sample iteration matrices for EQ and SK when applied to matrices C and D. Normalized iteration matrices, DAE, are printed with their row and column sums and deviations.

Algorithm EQ applied to matrix C

(total steps to convergence 34)

Normalized matrix at step 4

.5152426	.1703389	.0000000e+00	.6855814	-.3144186
.5126855e-01	1.694935	.1703389e-02	1.747907	.7479072
.0000000e+00	.5631393e-03	.5659481	.5665112	-.4334888

.5665112	1.865837	.5676515
-.4334888	.8658372	-.4323485

Normalized matrix at step 8

.8148247	.1101510	.0000000e+00	.9249758	-.7502425e-01
.8107809e-01	1.096044	.1999731e-02	1.179122	.1791216
.0000000e+00	.4907713e-03	.8954119	.8959027	-.1040973

.8959028	1.206686	.8974116
-.1040972	.2066857	-.1025884

Normalized matrix at step 12

.8862081	.9581436e-01	.0000000e+00	.9820224	-.1797760e-01
.8818100e-01	.9533885	.2018882e-02	1.043588	.4358840e-01
.0000000e+00	.4599243e-03	.9739290	.9743890	-.2561104e-01

.9743891	1.049663	.9759479
-.2561092e-01	.4966283e-01	-.2405208e-01

Normalized matrix at step 16

.9033740	.9236706e-01	.0000000e+00	.9957410	-.4258990e-02
.8988907e-01	.9190866	.2020057e-02	1.010996	.1099563e-01
.0000000e+00	.4517099e-03	.9928114	.9932631	-.6736875e-02

.9932630	1.011905	.9948315
-.6736994e-02	.1190531e-01	-.5168498e-02

Normalized matrix at step 20

.9076160	.9151518e-01	.0000000e+00	.9991312	-.8687973e-03
.9031117e-01	.9106101	.2020130e-02	1.002941	.2941370e-02
.0000000e+00	.4496311e-03	.9974777	.9979273	-.2072692e-02

.9979272	1.002575	.9994978
-.2072811e-02	.2574921e-02	-.5022287e-03

Normalized matrix at step 24

.9089828	.9151545e-01	.0000000e+00	1.000498	.4982948e-03
.9019836e-01	.9081077	.2014578e-02	1.000321	.3206730e-03
.0000000e+00	.4501961e-03	.9987310	.9991812	-.8187890e-03

.9991812	1.000073	1.000746
-.8188486e-03	.7343292e-04	.7456541e-03

Normalized matrix at step 28

.9091539	.9087843e-01	.0000000e+00	1.000032	.3242493e-04
.9086479e-01	.9082786	.9594032e-03	1.000103	.1028776e-03
.0000000e+00	.9456886e-03	.9989190	.9998647	-.1353025e-03

1.000019	1.000103	.9998784
.1871586e-04	.1027584e-03	-.1215935e-03

Normalized matrix at step 32

.9091351	.9086760e-01	.0000000e+00	1.000003	.2741814e-05
.9086480e-01	.9081892	.9594033e-03	1.000013	.1335144e-04
.0000000e+00	.9457083e-03	.9990382	.9999539	-.1609325e-04

.9999999	1.000003	.9999976
-.1192093e-06	.2503395e-05	-.2384186e-05

Normalized matrix at step 34

.9091351	.9086674e-01	.0000000e+00	1.000002	.1788139e-05
.9086479e-01	.9081804	.9594032e-03	1.000005	.4649162e-05
.0000000e+00	.9457082e-03	.9990476	.9999933	-.6735325e-05

.9999999	.9999928	1.000007
-.1192093e-06	-.7152557e-05	.7033348e-05

Algorithm SK applied to matrix C

(total steps to convergence 1899)

Normalized matrix at step 189

.9088035	.9044544e-01	.0000000e+00	.9992489	-.7510781e-03
.9119660e-01	.9076018	.4645516e-03	.9992629	-.7370710e-03
.0000000e+00	.1952808e-02	.9995354	1.001488	.1488209e-02

1.000000	1.0000000	.9999999
.1192093e-06	.0000000e+00	-.5960464e-07

Normalized matrix at step 378

.9089547	.9064316e-01	.0000000e+00	.9995979	-.4020929e-03
.9104527e-01	.9079254	.6338494e-03	.9996045	-.3955364e-03
.0000000e+00	.1431491e-02	.9993661	1.000798	.7976294e-03

1.0000000	1.0000000	.9999999
.0000000e+00	.0000000e+00	-.5960464e-07

Normalized matrix at step 567

.9090315	.9074026e-01	.0000000e+00	.9997717	-.2282858e-03
.9096844e-01	.9080542	.7527148e-03	.9997754	-.2246499e-03
.0000000e+00	.1205464e-02	.9992473	1.000453	.4527569e-03

.9999999	.9999999	1.0000000
-.1192093e-06	-.5960464e-07	.0000000e+00

Normalized matrix at step 756

.9090746	.9079351e-01	.0000000e+00	.9998681	-.1319051e-03
.9092539e-01	.9081141	.8306531e-03	.9998701	-.1298785e-03
.0000000e+00	.1092345e-02	.9991693	1.000262	.2616644e-03

1.0000000	.9999999	1.0000000
.0000000e+00	-.5960464e-07	.0000000e+00

Normalized matrix at step 945

.9090995	.9082385e-01	.0000000e+00	.9999233	-.7671118e-04
.9090054e-01	.9081445	.8794937e-03	.9999245	-.7551908e-04
.0000000e+00	.1031668e-02	.9991206	1.000152	.1522303e-03

1.0000000	1.0000000	1.000000
.0000000e+00	.0000000e+00	.1192093e-06

Normalized matrix at step 1134

.9091139	.9084138e-01	.0000000e+00	.9999552	-.4476309e-04
.9088606e-01	.9081606	.9092371e-03	.9999559	-.4410744e-04
.0000000e+00	.9979079e-03	.9990908	1.000089	.8869171e-04

.9999999	.9999999	1.0000000
-.5960464e-07	-.1192093e-06	.0000000e+00

Normalized matrix at step 1323

.9091224	.9085156e-01	.0000000e+00	.9999740	-.2604723e-04
.9087761e-01	.9081697	.9270366e-03	.9999743	-.2568960e-04
.0000000e+00	.9787399e-03	.9990729	1.000052	.5161762e-04

1.0000000	1.0000000	.9999999
.0000000e+00	.0000000e+00	-.5960464e-07

Normalized matrix at step 1512

.9091273	.9085749e-01	.0000000e+00	.9999848	-.1519918e-04
.9087269e-01	.9081748	.9375770e-03	.9999851	-.1490116e-04
.0000000e+00	.9677320e-03	.9990624	1.000030	.3015995e-04

1.0000000	1.0000000	1.0000000
.0000000e+00	.0000000e+00	.0000000e+00

Normalized matrix at step 1701

.9091302	.9086095e-01	.0000000e+00	.9999911	-.8881092e-05
.9086981e-01	.9081776	.9437799e-03	.9999912	-.8761883e-05
.0000000e+00	.9613687e-03	.9990562	1.000018	.1752377e-04

1.0000000	.9999999	.9999999
.0000000e+00	-.5960464e-07	-.5960464e-07

Normalized matrix at step 1890

.9091318	.9086296e-01	.0000000e+00	.9999948	-.5185604e-05
.9086813e-01	.9081794	.9474169e-03	.9999949	-.5066395e-05
.0000000e+00	.9576764e-03	.9990525	1.000010	.1025200e-04

.9999999	1.000000	.9999999
-.5960464e-07	.1192093e-06	-.5960464e-07

Algorithm EQ applied to matrix D

(total steps to convergence 40)

Normalized matrix at step 4

.5620255	.1772865e-04	.0000000e+00	.5620432	-.4379568
.5592362e-01	1.764067	.1772865e-04	1.820008	.8200079
.0000000e+00	.5592362e-01	.5620254	.6179491	-.3820509

.6179491	1.820008	.5620431
-.3820509	.8200079	-.4379569

Normalized matrix at step 8

.8912525	.2070484e-04	.0000000e+00	.8912731	-.1087269
.3629401e-01	.8431525	.2070484e-04	.8794672	-.1205328
.0000000e+00	.4809977e-01	1.181160	1.229260	.2292597

.9275465	.8912730	1.181181
-.7245350e-01	-.1087270	.1811807

Normalized matrix at step 12

.9768155	.2479174e-04	.0000000e+00	.9768403	-.2315974e-01
.3671759e-01	.9318984	.2094648e-04	.9686369	-.3136307e-01
.0000000e+00	.4491687e-01	1.009606	1.054523	.5452287e-01

1.013533	.9768401	1.009627
.1353300e-01	-.2315992e-01	.9626985e-02

Normalized matrix at step 16

.9915406	.2619333e-04	.0000000e+00	.9915668	-.8433223e-02
.3727109e-01	.9845832	.9836715e-03	1.022838	.2283800e-01
.0000000e+00	.9855222e-03	.9846095	.9855950	-.1440495e-01

1.028812	.9855949	.9855932
.2881169e-01	-.1440513e-01	-.1440680e-01

Normalized matrix at step 20

.9819534	.9902464e-03	.0000000e+00	.9829437	-.1705635e-01
.9878684e-03	.9962114	.9952888e-03	.9981946	-.1805425e-02
.0000000e+00	.1018789e-02	1.017845	1.018864	.1886404e-01
.9829413	.9982204	1.018841		
-.1705873e-01	-.1779556e-02	.1884055e-01		

Normalized matrix at step 24

.9998438	.9916134e-03	.0000000e+00	1.000835	.8354187e-03
.1005867e-02	.9975866	.9953964e-03	.9995879	-.4121065e-03
.0000000e+00	.1000776e-02	.9985785	.9995793	-.4206896e-03
1.000850	.9995790	.9995739		
.8496046e-03	-.4209876e-03	-.4260540e-03		

Normalized matrix at step 28

.9992091	.9919285e-03	.0000000e+00	1.000201	.2009869e-03
.1005228e-02	.9979036	.9953965e-03	.9999042	-.9578466e-04
.0000000e+00	.1001411e-02	.9988956	.9998971	-.1029372e-03
1.000214	.9998969	.9998910		
.2143383e-03	-.1030564e-03	-.1089573e-03		

Normalized matrix at step 32

.9990506	.9920071e-03	.0000000e+00	1.000043	.4267693e-04
.1005069e-02	.9979827	.9953965e-03	.9999832	-.1680851e-04
.0000000e+00	.1001570e-02	.9989749	.9999765	-.2348423e-04
1.000056	.9999763	.9999703		
.5567074e-04	-.2366304e-04	-.2968311e-04		

Normalized matrix at step 36

.9990111	.9920268e-03	.0000000e+00	1.000003	.3099442e-05
.1005029e-02	.9980026	.9953966e-03	1.000003	.2980232e-05
.0000000e+00	.1001610e-02	.9989947	.9999963	-.3695488e-05

1.000016	.9999962	.9999901
.1609325e-04	-.3814697e-05	-.9894371e-05

Normalized matrix at step 38

.9990011	.9920268e-03	.0000000e+00	.9999931	-.6854534e-05
.1005019e-02	.9980025	.9953966e-03	1.000003	.2861023e-05
.0000000e+00	.1001620e-02	.9990047	1.000006	.6318092e-05

1.000006	.9999961	1.0000000
.6198883e-05	-.3874302e-05	.0000000e+00

Algorithm SK applied to matrix D

(total steps to convergence 2983)

Normalized matrix at step 298

.9967212	.3029083e-03	.0000000e+00	.9970241	-.2975881e-02
.3278699e-02	.9964123	.3032447e-03	.9999942	-.5781651e-05
.0000000e+00	.3284839e-02	.9996968	1.002982	.2981663e-02
.9999999	1.0000000	1.0000000		
-.1192093e-06	.0000000e+00	.0000000e+00		

Normalized matrix at step 596

.9981694	.5439625e-03	.0000000e+00	.9987133	-.1286685e-02
.1830637e-02	.9976242	.5442962e-03	.9999991	-.8940697e-06
.0000000e+00	.1831872e-02	.9994557	1.001288	.1287580e-02
1.0000000	1.0000000	1.0000000		
.0000000e+00	.0000000e+00	.0000000e+00		

Normalized matrix at step 894

.9986148	.7194218e-03	.0000000e+00	.9993342	-.6657839e-03
.1385158e-02	.9978949	.7196584e-03	.9999997	-.2980232e-06
.0000000e+00	.1385625e-02	.9992803	1.000666	.6659031e-03
.9999999	.9999999	1.0000000		
-.5960464e-07	-.5960464e-07	.0000000e+00		

Normalized matrix at step 1192

.9988053	.8343542e-03	.0000000e+00	.9996397	-.3603101e-03
.1194671e-02	.9979708	.8345041e-03	.9999999	-.5960464e-07
.0000000e+00	.1194887e-02	.9991654	1.000360	.3602505e-03
1.0000000	1.0000000	.9999999		
.0000000e+00	.0000000e+00	-.5960464e-07		

Normalized matrix at step 1490

.9988979	.9045403e-03	.0000000e+00	.9998024	-.1975894e-03
.1102099e-02	.9979932	.9046296e-03	.9999999	-.1192093e-06
.0000000e+00	.1102209e-02	.9990954	1.000198	.1975298e-03

.9999999	.9999999	1.0000000
-.5960464e-07	-.5960464e-07	.0000000e+00

Normalized matrix at step 1788

.9989457	.9455816e-03	.0000000e+00	.9998913	-.1087189e-03
.1054322e-02	.9979999	.9456330e-03	.9999999	-.1192093e-06
.0000000e+00	.1054380e-02	.9990544	1.000109	.1087189e-03

1.0000000	.9999999	1.0000000
.0000000e+00	-.1192093e-06	.0000000e+00

Normalized matrix at step 2086

.9989711	.9689744e-03	.0000000e+00	.9999401	-.5990267e-04
.1028897e-02	.9980021	.9690035e-03	1.0000000	.0000000e+00
.0000000e+00	.1028928e-02	.9990309	1.000060	.5984306e-04

1.0000000	1.0000000	.9999999
.0000000e+00	.0000000e+00	-.1192093e-06

Normalized matrix at step 2384

.9989848	.9821138e-03	.0000000e+00	.9999669	-.3308058e-04
.1015147e-02	.9980027	.9821299e-03	.9999999	-.5960464e-07
.0000000e+00	.1015164e-02	.9990178	1.000033	.3302097e-04

.9999999	1.0000000	.9999999
-.5960464e-07	.0000000e+00	-.5960464e-07

```
Normalized matrix at step      2682

     .9989924        .9894333e-03    .0000000e+00       .9999819      -.1811981e-04
     .1007645e-02    .9980029        .9894422e-03      1.0000000       .0000000e+00
     .0000000e+00    .1007654e-02    .9990106          1.000018        .1823902e-04

    1.000000        1.0000000       1.0000000

     .1192093e-06    .0000000e+00    .0000000e+00

Normalized matrix at step      2980

     .9989966        .9934918e-03    .0000000e+00       .9999900      -.9953976e-05
     .1003533e-02    .9980029        .9934969e-03       .9999999      -.5960464e-07
     .0000000e+00    .1003538e-02    .9990065          1.000010        .1001358e-04

    1.000000        1.0000000        .9999999

     .1192093e-06    .0000000e+00   -.5960464e-07
```

This report was done with support from the Center
for Pure and Applied Mathematics. Any conclusions
or opinions expressed in this report represent solely
those of the author(s) and not necessarily those of
The Regents of the University of California, the
Center for Pure and Applied Mathematics or the
Department of Mathematics.

DATE

ILMED

-8